CONOPS PURPOSE

- Communicate the user's needs for and expectations of the prosed system to the developers,

- communicate the developers understanding of the users' need and how the system shall operate to fulfill those needs,

- build consensus among several user groups, among several buyer organizations, and/or among several developers.

Paraphrased from IEEE Standard 1362

AUDIENCE

- Users might read it to determine whether their needs and desires have been correctly specified... or to verify the developers understanding of their needs.

- Buyers might read it to acquire knowledge of the user's needs and or developer's understanding of those needs.

- Developers will typically use the ConOps document as a basis for system development activities, and to familiarize now team members with the problem domain and the system to which the ConOps applies.

ConOps Sections

1. Scope

2. Reference documents

3. Current system or situation

4. Justification for and nature of changes

5. Concepts for the proposed system

6. Operational scenarios

7. Summary of impacts

8. Analysis of the proposed system

1. Scope

The standard specifies the following subsections:
- 1.1 Identification
- 1.2 Document purpose and audience
- 1.3 System overview

The last "briefly states the purpose of the proposed system" and "describes the general nature of the system, and identifies the project sponsors, support agencies, certifiers or certifying bodies, and the operating centers or sites that will run the system."
"A graphical overview of the system is strongly recommended. This can be in the form of a context diagram, a top level object diagram" etc.

GRAPHICAL OVERVIEW
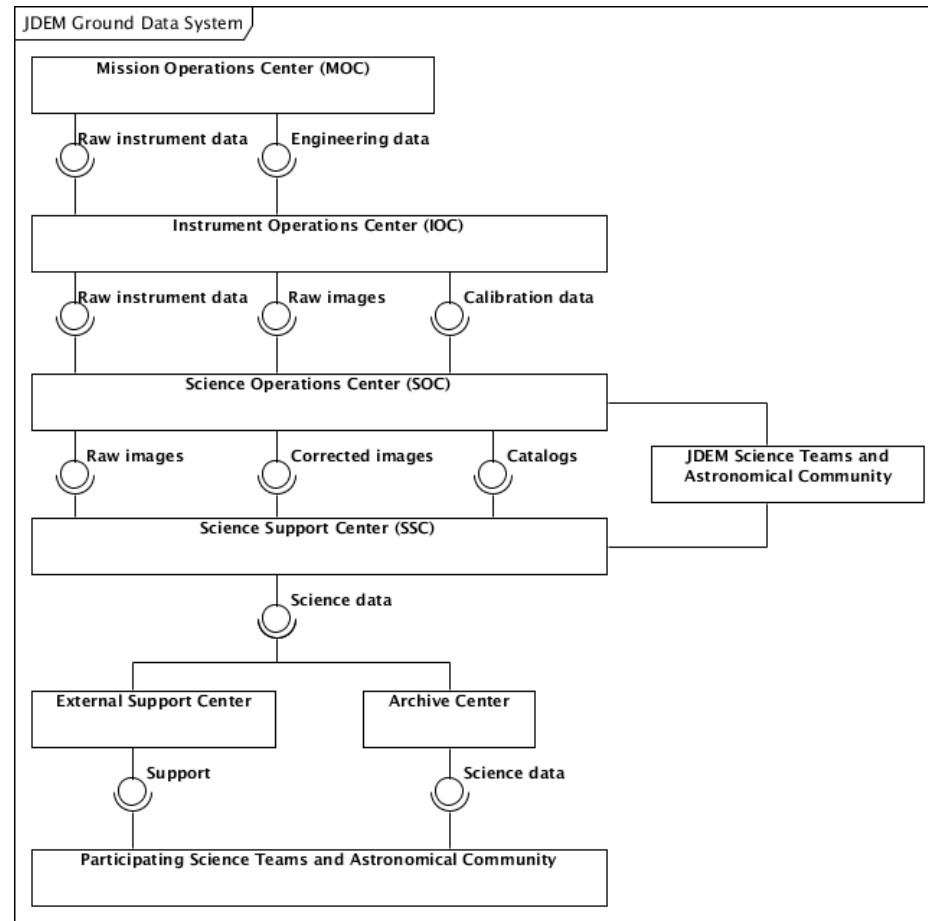
3. CURRENT SYSTEM OR SITUATION

- 3.1 Background, objectives, and scope

- 3.2 Operational policies and constraints

- 3.3 Description of the current system or situation

- 3.4 Modes of operation

- 3.5 User classes and other involved personnel

- 3.6 Support environment

3.3 DESCRIPTION OF THE CURRENT SYSTEM OR SITUATION

- Expertise and experience (SDSS, ...)
- Processing infrastructure (Fermigrid)
- Data management infrastructure (dCache, robots)
- Workflow management (DAGMan, sdssdp, GGG)
- Bookkeeping and provenance (SDSS, lqcd)
- Quality control and monitoring (NGOP?)
- Applications (SDSS pipeline, IRAF, etc.)

4 Justification for and nature of changes

- 4.1 Justification of changes
- 4.2 Description of desired changes
- 4.3 Priorities among changes
- 4.4 Changes considered but not included

5. CONCEPTS FOR THE PROPOSED SYSTEM

- 5.1 Background, objectives, and scope

- 5.2 Operational policies and constraints

- 5.3 Description of the proposed system

- 5.4 Modes of operation

- 5.5 User classes and other involved personnel

- 5.6 Support environment

5.3 DESCRIPTION OF THE PROPOSED SYSTEM

- Overview

- Requirements

- High-level design for a demonstration data processing workflow

- A demonstration data processing workflow

- A workflow engine

- Provenance tools

- Data storage

- Monitoring and QC tools

5.3.1 Overview

- a set of requirements for a demonstration data processing workflow

- a high level design for a demonstration data processing workflow, including a data model

- a data processing workflow consisting of dummy and prototype participants

- prototype implementations of participants needed for processing JDEM spectroscopic data

- a workflow engine that executes the demonstration workflow

- tools for monitoring the state of the demonstration workflow

- tools for recording provenance of demonstration workflow data products

- one or more data storage elements that produce and consume data read and generated by the workflow

5.3.2 REQUIREMENTS

Requirements will be collected and tracked in DOORS.

5.3.3 HIGH-LEVEL DESIGN FOR A DEMONSTRATION DATA PROCESSING WORKFLOW

- a textual overview of the workflow,

- a list of workflow participants, with brief high level descriptions of their required actions;

- one or more UML 2 Activity diagram showing relationships between the actions performed by the workflow participants;

- a list of data elements, describing the data which to be transferred between processing elements and into and out of data stores;

- specifications for workflow participant interfaces,

- a data model for the storage of data in data stores.

5.3.4 THE DEMONSTRATION DATA PROCESSING WORKFLOW

The demonstration data processing workflow will consist of a set of workflow participants, sample data, and any configuration and interface code needed for integration with other elements of the infrastructure, including the workflow engine, monitoring and QC tools, and data storage.

Participants required for processing spectroscopic data will be prototype implementations, and collectively be capable of calculating redshifts from simulated data. Other participants will be stubs only, accepting and producing data on the the specified interfaces but executing no useful calculations.

5.3.5 A WORKFLOW ENGINE

The demonstration workflow will be executed using the workflow engine being developed for Lattice QCD.

5.3.6 PROVENANCE

Provenance of data and the execution history of participants well be stored using the configuration management system being developed for Lattice QCD.

5.3.7 DATA STORAGE

Database and file storage adequate for supplying data to the workflow and storing data produced by it will be developed.

MONITORING AND QC TOOLS

A mechanism for monitoring the status of the workflow and reporting data metrics will be developed.

6. ANALYSIS OF THE PROPOSED SYSTEM

- 6.1 Summary of improvements

- 6.2 Disadvantages and limitations

- 6.3 Alternatives and trade-offs considered

REMAINING ISSUES

- The ConOps only barely touches on the simulation of spectroscopic data. We need to answer some questions on it:

    - How much of the sky/how many objects?

    - How many detectors?

    - What aspects should be simulated?

- Better define the scope of the prototype workflow

    - Does it include, for example, generation of observing plans? Application of calibrations and nothing else?

Document Flow

- As a whole, the document doesn't flow well.

- For a single subsystem (e. g. the the spectroscopic pipeline) the current system → needed changes → new system concepts organization might flow well. For the set, it doesn't seem to.

- Should we make this top-level document shorter, just listing the subsystems but not describing them, and then write separate short ConOps documents for each subsystem?

LEVEL OF DETAIL

- The current draft describes some systems, such as the spectroscopic pipeline, at about the level of detail I want right now. Lots of refinement and editing is needed, but not much expansion, yet.

- Other systems, such as the workflow engine, need more detail.

ConOps Purpose

- Communicate the user's needs for and expectations of the prosed system to the developers,

- communicate the developers understanding of the users' need and how the system shall operate to fulfill those needs,

- build consensus among several user groups, among several buyer organizations, and/or among several developers.

Paraphrased from IEEE Standard 1362

1

It's sort of a cross between a proposal and very high level requirements.

I personally want this so I can get a clear understanding of what my deliverables are.

AUDIENCE

- Users might read it to determine whether their needs and desires have been correctly specified... or to verify the developers understanding of their needs.

- Buyers might read it to acquire knowledge of the user's needs and or developer's understanding of those needs.

- Developers will typically use the ConOps document as a basis for system development activities, and to familiarize now team members with the problem domain and the system to which the ConOps applies.

Quoted from IEEE Standard 1362

2

In our context, I think Deb qualifies as a buyer, and we ourselves are both developers and users. Science teams might also be considered users.

ConOps Sections

1. Scope

2. Reference documents

3. Current system or situation

4. Justification for and nature of changes

5. Concepts for the proposed system

6. Operational scenarios

7. Summary of impacts

8. Analysis of the proposed system

3

Guided by IEEE 1362

Following the template makes you think about things systematically, but the flow of the document stinks, not every section is useful for any given project.

1. SCOPE

The standard specifies the following subsections:
- 1.1 Identification
- 1.2 Document purpose and audience
- 1.3 System overview

The last "briefly states the purpose of the proposed system" and "describes the general nature of the system, and identifies the project sponsors, support agencies, certifiers or certifying bodies, and the operating centers or sites that will run the system."
"A graphical overview of the system is strongly recommended. This can be in the form of a context diagram, a top level object diagram" etc.

The first two are trivial.

Should work elements from the "purpose" document into this section. The stakeholders section of the purpose document should probably just be copied into the ConOps almost word for word.

The purpose document lists two purposes- to strengthen Fermilab's case for hosting the SOC, and to reduce the cost and risk of hosting the SOC. The ConOps draft as it is now only mentions the second. Should the first be worked in as well? Is there anything that would be motivated by the first and not the second? Documentation perhaps?

GRAPHICAL OVERVIEW
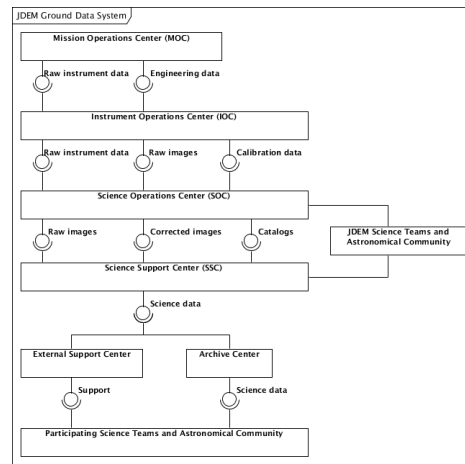
This is a UML component diagram created in VP-UML. The "lollipop" notation seems useless and messy to me.

3. Current System or Situation

- 3.1 Background, objectives, and scope
- 3.2 Operational policies and constraints
- 3.3 Description of the current system or situation
- 3.4 Modes of operation
- 3.5 User classes and other involved personnel
- 3.6 Support environment

6

Except for 3.1 and 3.3, these elements are trivial.

It is unclear to me what to put for 3.1, or if it is useful. What I have there now is not, I think, appropriate.

Overall, this section may not be very valuable when the audience is someone at Fermilab, but may be a very good thing for others in the collaboration (e. g. Deb Aragwal). Having this here should help us use this document to explain effort estimates, etc.

3.3 Description of the current system or situation

- Expertise and experience (SDSS, ...)
- Processing infrastructure (Fermigrid)
- Data management infrastructure (dCache, robots)
- Workflow management (DAGMan, sdssdp, GGG)
- Bookkeeping and provenance (SDSS, lqcd)
- Quality control and monitoring (NGOP?)
- Applications (SDSS pipeline, IRAF, etc.)

7

Under "Expertise and experience," I cover what the EAG folk (Steve and I) have done at least a little bit; what the CET has done still needs to be added.

In it's current state, the descriptions of the various existing systems are very brief, only a sentence or two. To be useful, enough detail should be added to give collaborators and management outside of Fermilab a good idea of what Fermilab can "just do," what needs a little development work, and what needs a lot. Perhaps this is analogous to a resume?

4 JUSTIFICATION FOR AND NATURE OF CHANGES

- 4.1 Justification of changes
- 4.2 Description of desired changes
- 4.3 Priorities among changes
- 4.4 Changes considered but not included

8

This section ends up being redundant with or at least strongly implied by much of 3.3, although this might be because some parts of section 3 should be moved here.

Currently it is only about a page long, and doesn't say much.

5. CONCEPTS FOR THE PROPOSED SYSTEM

- 5.1 Background, objectives, and scope
- 5.2 Operational policies and constraints
- 5.3 Description of the proposed system
- 5.4 Modes of operation
- 5.5 User classes and other involved personnel
- 5.6 Support environment

9

Again, 5.1 appears to be redundant with other sections.

Section 5.3 is the heart of the section, and is effectively a list of deliverables.

5.3 DESCRIPTION OF THE PROPOSED SYSTEM

- Overview

- Requirements

- High-level design for a demonstration data processing workflow

- A demonstration data processing workflow

- A workflow engine

- Provenance tools

- Data storage

- Monitoring and QC tools

These sections are effectively lists of deliverables.

5.3.1 OVERVIEW

- a set of requirements for a demonstration data processing workflow

- a high level design for a demonstration data processing workflow, including a data model

- a data processing workflow consisting of dummy and prototype participants

- prototype implementations of participants needed for processing JDEM spectroscopic data

- a workflow engine that executes the demonstration workflow

- tools for monitoring the state of the demonstration workflow

- tools for recording provenance of demonstration workflow data products

- one or more data storage elements that produce and consume data read and generated by the workflow

11

This slide is just a copy of the current section of the document, except for a short sentence at the top to introduce the list.

5.3.2 Requirements

Requirements will be collected and tracked in DOORS.

That's all I have.

5.3.3 High-level design for a demonstration data processing workflow

- a textual overview of the workflow,

- a list of workflow participants, with brief high level descriptions of their required actions;

- one or more UML 2 Activity diagram showing relationships between the actions performed by the workflow participants;

- a list of data elements, describing the data which to be transferred between processing elements and into and out of data stores;

- specifications for workflow participant interfaces,

- a data model for the storage of data in data stores.

This is effectively a list a deliverables.

5.3.4 THE DEMONSTRATION DATA PROCESSING WORKFLOW

The demonstration data processing workflow will consist of a set of workflow participants, sample data, and any configuration and interface code needed for integration with other elements of the infrastructure, including the workflow engine, monitoring and QC tools, and data storage.

Participants required for processing spectroscopic data will be prototype implementations, and collectively be capable of calculating redshifts from simulated data. Other participants will be stubs only, accepting and producing data on the the specified interfaces but executing no useful calculations.

5.3.5 A workflow engine

The demonstration workflow will be executed using the workflow engine being developed for Lattice QCD.

5.3.6 PROVENANCE

Provenance of data and the execution history of participants well be stored using the configuration management system being developed for Lattice QCD.

5.3.7 DATA STORAGE

Database and file storage adequate for supplying data to the workflow and storing data produced by it will be developed.

Monitoring and QC tools

A mechanism for monitoring the status of the workflow and reporting data metrics will be developed.

This is QC, not QA. From wikipedia: "quality control emphasizes testing of products to uncover defects, and reporting to management who make the decision to allow or deny the release, whereas quality assurance attempts to improve and stabilize production, and associated processes, to avoid, or at least minimize, issues that led to the defects in the first place." Textbooks I have seen say something similar, but less concisely.

6. ANALYSIS OF THE PROPOSED SYSTEM

- 6.1 Summary of improvements
- 6.2 Disadvantages and limitations
- 6.3 Alternatives and trade-offs considered

I have not written anything here yet.

Remaining Issues

- The ConOps only barely touches on the simulation of spectroscopic data. We need to answer some questions on it:
    - How much of the sky/how many objects?
    - How many detectors?
    - What aspects should be simulated?
- Better define the scope of the prototype workflow
    - Does it include, for example, generation of observing plans? Application of calibrations and nothing else?

20

Document Flow

- As a whole, the document doesn't flow well.

- For a single subsystem (e. g. the the spectroscopic pipeline) the current system $\rightarrow$ needed changes $\rightarrow$ new system concepts organization might flow well. For the set, it doesn't seem to.

- Should we make this top-level document shorter, just listing the subsystems but not describing them, and then write separate short ConOps documents for each subsystem?

LEVEL OF DETAIL

- The current draft describes some systems, such as the spectroscopic pipeline, at about the level of detail I want right now. Lots of refinement and editing is needed, but not much expansion, yet.

- Other systems, such as the workflow engine, need more detail.

22